Description Generation of Abnormal Densities found in Radiographs

Alicia Abella, Ph.D.,
AT&T Bell Laboratories, Murray Hill, NJ
John R. Kender, Ph.D.,
Computer Science Department, Columbia University, New York, NY
Justin Starren, M.D.,
Department of Medical Informatics, Columbia University, New York, NY

Abstract

In this paper we present a system for describing renal stones found in radiographs. The system generates descriptions that adhere to those generated by radiologists. The descriptions are formulated by discovering the spatial relationships that exist between the major organs and the renal stones. The system consists of three major components. The first is the image processing component which is responsible for locating the stone. The second component is the inference network minimization component which determines which spatial relationships, of all those that exist between the stone and the organs, is the most descriptive. The third component is the natural language generation component which is responsible for translating the spatial relationships into appropriate medical terminology. We will illustrate all these components on several examples.

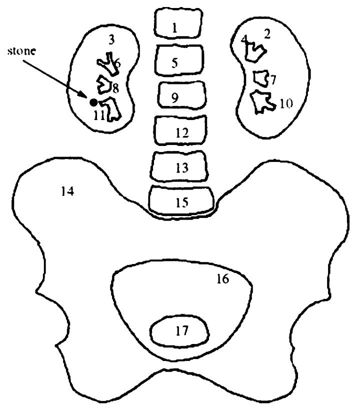

Figure 1: The model of the urinary system. (3 right kidney), (2 left kidney), (4,6,10,11 calyx), (8,7 middle-calyx), (5 L1), (9 L2), (12 L3), (13 L4), (15 L5), (14 pelvis), (16 inner pelvis) (17 bladder).

Introduction

A considerable amount of work has been done on the image processing aspects of locating abnormalities in radiographs. This work goes beyond simply locating the abnormality to generating descriptions of the abnormalities. Specifically the goal of the project presented in this paper is to describe renal stones found in radiographs. The issue is to first determine if a stone is present in the radiograph, and if so to identify its location for possible treatment. To describe the location involves discovering the spatial relationships of the stone to reference objects in the image. Radiologists categorize stones according to their locations on the X-ray. The stones are described by observing the spatial relationship of the stones to other parts of the radiograph. Those parts include the spinal cord, the various lumbar bodies of the spinal cord, the bladder, the kidney, the calyx, and a few others that are not usually visible on the X-ray. The ones we use are shown in figure 1. Our system is unique in its ability to integrate image processing and natural language processing for the task of describing renal stones found in radiographs. Most research efforts have focused on either of the two. Somewhat along the lines of our work is the work by [Fox and Walker, 1989] who believe that a useful role of computers in medicine is for imaging systems to be combined with methods for interpreting clinical data.

More along the lines of traditional medical imaging systems can be found in [Kobashi and Shapiro, 1992]. They describe a knowledge-based recognition system that utilizes knowledge of anatomy and CT (computerized tomography) imaging for organ identification of the abdomen. A method for automatically detecting boundaries of brain tumors is given in [Lu *et al.*, 1992]. A similar paper is [Selfridge and Prewitt, 1981] that describes two boundary-delineating algorithms for detecting kidneys in tomographic images. Most medical image processing involves standard techniques like those found in [Wechsler and Sklansky, 1977]. In this paper the authors describe a system for finding the rib cage in chest radiographs.

Image Processing

Before we can describe a potential renal stone, we must be able to find it in the image. This work does not develop any new algorithms for image processing of X-ray images, instead we use existing algorithms. Our contribution is in the area of description generation. The image processing component of the system needs to accomplish two separate yet interdependent goals,

0195-4210/95/$5.00 © 1995 AMIA, Inc.

the first is to register the image with the model of the urinary system shown in figure 1 and the second is to find the stone once the image has been registered. This is necessary because each image is slightly distorted. The reasons for the distortions are due to differences in anatomy and radiologic techniques.

Accurately describing the location of the stone will depend on how well the system was able to register the image and locate the stone.

Generating the binary image. The image processing module first converts an X-ray image into a black and white (binary) image in order to locate the spinal cord and pelvis. The spinal cord and pelvis are the two landmarks that we extract from the X-ray image in order to register it with the model.

Edge Detection. Once we have created the binary image we need to find the edges. The edges are small regions in an image that have a rapid change in image intensity. To find the edges we use a 5x5 edge operator, patterned after the Sobel 3x3, as a discrete approximation for the partial derivatives that measure the gradient [Horn, 1990].

Registering the image. In this step the image processing module needs to find a transformation that maps the edge image into the model. The goal is to bring the image as close as possible to the model image. The system looks for a six parameter affine coordinate transformation that accomplishes this and applies it to the image. We are then ready to use this transformed image to locate the stones.

Locating the stones.
To locate the stones we use the technique sketched out in [Kimme *et al.*, February 1975] for circle finding. Since a majority of the stones are circular in nature we may use this technique.

Once the stone is found it is superimposed on the model and we are ready to determine its position relative to the various reference objects of figure 1.

Inference Network Minimization

In this section we will describe how the system chooses which spatial relationships best describe the position of the stone relative to the reference objects from the model. The spatial relationships are represented by computational models of a set of spatial prepositions that include: {*inside*, *near*, *left*, *right*, *above*, *below*, *between*}. A more detailed exposition about the computational models can be found in [Abella, 1995]. The system uses these computational models to determine if two objects (the stone and an organ or bony structure) are in a particular spatial relationship (e.g. `The stone is` *inside* `the right kidney.`).

Many of the spatial relationships that the system will find as being true are redundant in the final description, this is why the system is equipped with an algorithm to eliminate unnecessary relationships between the figure object (*the stone*) and the reference objects (*the kidney*, *the bladder* etc). The spatial relationships are all expressed through prepositions. The

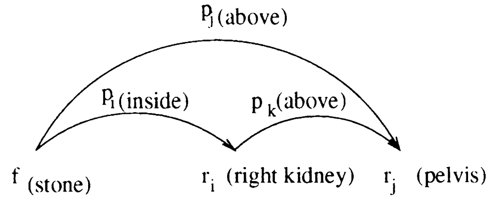

Figure 2: The inference rule: if a preposition p_k can be found that together with p_i implies preposition p_j then the relationship between reference object r_j and the figure object f may be eliminated from the final description.

system eliminates those prepositions that are redundant. When eliminating these prepositions we need to look at a pair of reference objects r_i and r_j and the prepositions p_i and p_j that are true for (f, r_i) and (f, r_j) where f is the stone. If we can find a preposition p_k that relates r_i and r_j such that the following condition holds

$$(\forall r_i, f, r_j)(p_k(r_i, r_j) \wedge p_i(f, r_i) \Longrightarrow p_j(f, r_j)) \qquad (1)$$

then preposition p_j is redundant and we may eliminate it from the final description. Figure 2 graphically illustrates this condition. For example, suppose r_i is the right kidney, r_j is the pelvis, and the stone is *inside* the right kidney as illustrated in figure 2. In this case, p_i is *inside* and p_j is *above*. Choosing p_k = *above* allows us to eliminate the fact that the stone is *above* the pelvis because the radiologist knows that the right kidney is *above* the pelvis and it is already known that the stone is *inside* the right kidney. Thus, it is not necessary to say that the stone is also *above* the pelvis.

The graph in figure 3 illustrates the inferences we use to eliminate as many relationships as possible[1]. Note that these inferences are independent of the domain; they are provable based on the computational models of the spatial prepositions. Each node represents a preposition that relates a reference object and a figure object. An edge between two nodes is labeled by the relationship between two reference objects that needs to hold in order to eliminate the preposition this edge is pointing to.

In determining if we can eliminate a spatial relationship we follow the edges in the graph in the following manner. We begin at the node p_i that describes the relationship between figure object f and reference object r_i. If there exists an edge out of p_i that describes the relationship between r_i and r_j we follow that link to the next node p_j. If p_j describes the spatial relationship between f and r_j then we may eliminate it from the final description, because p_i and p_k imply p_j. For example let us follow the edge from *near* back to itself. We begin at the node labeled *near* if we know that f is *near* r_i. We follow the edge out of the node

[1] The prepositions preceded by the word restricted have a more constrained definition than the prepositions that are not restricted. This was necessary in order to maintain the provability of the inferences. Since this issue is not the topic of this paper, we will not elaborate further, but more details can be found in [Abella, 1995]

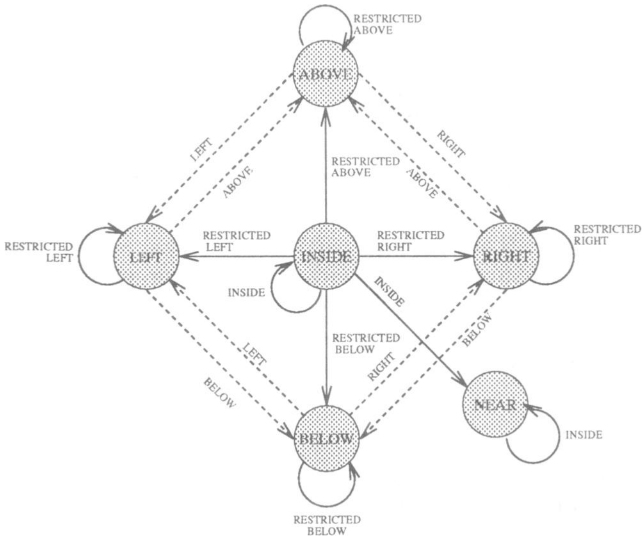

Figure 3: The inferences used to eliminate relationships from the final description. Nodes represent prepositions p for object pairs (f, r_i) and edge labels represent prepositions p for object pairs (r_i, r_j).

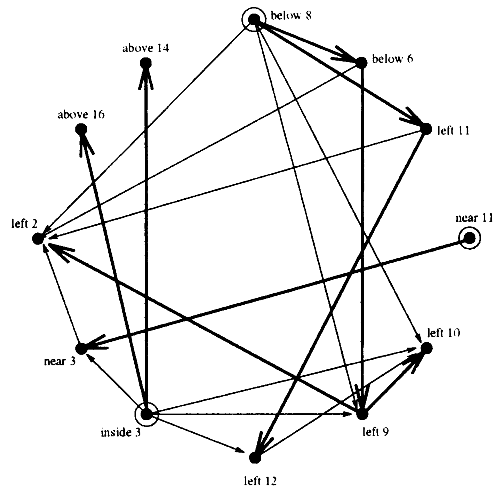

Figure 4: The graph G and its spanning forest (thick edges). The roots of the spanning forest are circled.

if r_i is *inside* r_j. If it is we may eliminate the fact that f is *near* r_j.

The dashed links in figure 3 represent weak links. Weak links do not satisfy condition 1, but a somewhat weaker condition which is that the *complement* of the preposition p_j is *not* satisfied:

$$(\forall r_i, f, r_j)(p_k(r_i, r_j) \wedge p_i(f, r_i) \Longrightarrow \neg - p_j(f, r_j))$$

where $-p_j$ stands for the complement of p_j. Not all prepositions have complements. Obvious complement pairs are (*left*, *right*) and (*above*, *below*).

Encoding inferences using spanning trees

This section describes how to extract the minimal set of necessary descriptions based on the graph of figure 3. We begin with a set of all possible descriptions of a figure object with respect to all the reference objects. These descriptions can be thought of as a pair (p, r) where p is a preposition and r is a reference object, such that $p(f, r) = 1$, which means that preposition p describes the relationship between figure object f and reference object r. Not all of these descriptions are required to describe the figure object. The graph of figure 3 will enable us to detect the redundant descriptions. This graph defines a ternary relation $T(p_1, p_2, p_3)$. We say that prepositions p_1, p_2, p_3 are in relation T if the graph in figure 3 contains an edge pointing from p_1 to p_2 and labeled p_3. For example, $T(\mathit{inside}, \mathit{near}, \mathit{inside}) = 1$. Using this relation we will define a directed graph G whose nodes are all possible descriptions (p, r) of a figure object. Descriptions (p_1, r_1) and (p_2, r_2) are connected by an edge from (p_1, r_1) to (p_2, r_2) (denoted $(p_1, r_1) \rightarrow (p_2, r_2)$) if a preposition p_3 exists such that r_1 and r_2 are related by p_3 and p_1, p_2, and p_3 are related by T:

$$(p_1, r_1) \rightarrow (p_2, r_2) \Leftrightarrow (\exists p_3)(p_3(r_1, r_2) \wedge T(p_1, p_2, p_3))$$

For example if f is *inside* r_1 and *near* r_2 and if r_1 is *inside* r_2 then $(\mathit{inside}, r_1) \rightarrow (\mathit{near}, r_2)$ since $T(\mathit{inside}, \mathit{near}, \mathit{inside})$.

If a node has an ancestor in G then the description conveyed by that ancestor would render the description conveyed by the node redundant. Therefore the minimal description consists of the roots of the spanning forest of G. In an acyclic graph the roots of a spanning forest cannot be inferred from any other nodes, however all other nodes can be inferred by them. If the graph contains cycles then the spanning forest is not unique and we must apply an ordering on the nodes. The ordering is based on the importance of the node. For example, a node that describes a relationship that involves the kidney is more important than one that involves the pelvis.

Inference Example

We will present an example of the inference network minimization using the image in figure 1. This image represents a model of the urinary system with a stone found in the right[2] kidney. Each object in the model (both organs and bony structure, e.g. kidney and spinal cord) are numbered to simplify the presentation. The first step in retrieving the minimal set of descriptions is to compute all the spatial relation of the stone to all the objects in the model. The result is a list of stone descriptions with respect to the different objects. This list is

```
((below 8) (below 6) (left 11) (near 11)
(left 10) (left 9) (left 12) (inside 3)
(near 3) (left 2) (above 16) (above 14))
```

From these descriptions and the graph of figure 3 we form the implication graph shown in figure 4. For example, an edge from the node labeled **below 8** to the node labeled **below 6** exists because the stone is *below* 8 and *below* 6 and according to the graph this edge may be drawn if 8 is *below* 6 which it is. In other words, the fact that the stone is *below* 8 and that 8

[2]The right kidney appears left in the image.

is *below* 6 makes saying that the stone is *below* 6 unnecessary. The other edges are generated in this same manner. This figure also shows the spanning forest of this graph whose roots are the minimal description for the stone: `((below 8) (near 11) (inside 3))`. We will see in the next section how this minimal description translates to the sentence **`The right lower quadrant contains a density which may represent a stone in the lower pole calyx.`**

Natural Language Generation

In this section we will discuss how the spatial relations of the previous section are translated into the proper medical terminology. Each preposition or combination of prepositions can potentially map into a particular description of a density.

The inference network minimization algorithm supplies all the spatial relations found to be necessary and sufficient to describe the stone. It is the job of the language generation module to compose the appropriate input to the natural language generator so that it may produce a meaningful sentence similar to the types of sentences that could be found in actual radiology reports. The language generation module is an embryo of a rule-based system for translating spatial relations into proper medical terminology. The rules were defined by using the radiology reports associated with the images. The natural language generator we used is called FUF (Functional Unification Formalism) [Elhadad, 1993] and it is responsible for generating the final English sentences.

An example of input to the language generation module is
`((inside right-kidney) (near calyx) (above middle-calyx))`
The language generation module then expands this input and creates the input needed by the natural language generator. For this particular example the fact that the stone is *above* the middle-calyx signals the language generation module that the stone is in the upper portion of the right kidney. The language generation module translates this into the proper medical term *upper pole*.

The pair **`(near calyx)`** causes the language generation module to generate the semantic input that will produce the phrase *upper pole calyx* or *lower pole calyx* depending on whether the stone is *above* or *below* the middle calyx.

The final sentence produced by the natural language generator is
`The right upper quadrant contains a density which probably represents a stone in the upper pole calyx.`

Clinical Examples

We tested our system on five radiographs, each exhibiting a stone in a different location. The system was successful in locating and describing the stone in four out of the five cases. In the fifth case the stone was too small to be discriminated from noise in the image.

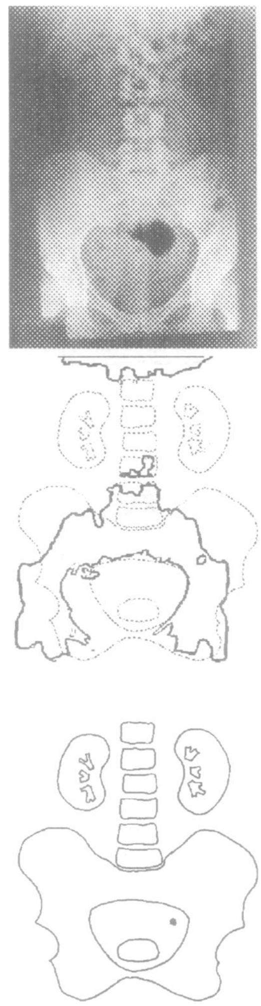

Figure 5: Original X-ray (top); Transformed edge image superimposed on the model (middle); Stone superimposed on the model (bottom)

Example 1: Distal Left Ureteral Stone

Figure 5 shows the original X-ray, the transformed edge image superimposed on the model, and the stone that was found superimposed on the model. After applying the inference network minimization technique the spatial relations that resulted were **`((inside inner-pelvis) (right inner-pelvis))`**. This was then translated by the language generation preprocessor and the proper semantic input was sent to the language generator to produce the following sentence:
`A density is seen in the distal left ureter.`

Example 2: Renal Stone

Figure 6 shows the original X-ray, the transformed edge image superimposed on the model, and the stone that was found superimposed on the model. The spatial relations found for this example were **`((inside right-kidney) (left calyx))`**. The sentence that the natural language generator produced for this example was:
`The right kidney contains a density which`

Figure 6: Original X-ray (top); Transformed edge image superimposed on the model (middle); Stone superimposed on the model (bottom)

may represent a right renal stone.

Example 3: Calyceal Stone

For the third example the system found the following spatial relations: **(inside right-kidney) (near calyx) (below middle-calyx)**. The sentence the system generated was:
The right lower quadrant contains a density which may represent a stone in the lower pole calyx.

Example 4: Mid-Ureteral Stone

For the fourth example the system found the following spatial relations: **((left L4) (near right-kidney) (inside pelvis))**. The sentence the system generated was
A density is seen at the level of L4 on the right which may represent a stone in the right mid-ureter.

Conclusion

This paper covered the material necessary to gain an understanding of what is involved in the image processing and language generation components of a system that can generate medically sound descriptions of stones found in radiographs. The descriptions are medically sound because they are formulated by the rule-based system whose rules were defined using actual radiology reports.

We have included in the language generation module the capability of generating the most commonly occurring stone descriptions. Further work is needed to generate less common descriptions of stones as well as different kidney diseases such as tumors or phleboliths. This will require collection and testing of more images.

Acknowledgments

The authors wish to thank Dr. Jeff Newhouse for diligently collecting and evaluating the radiographs that went into this work. This publication was supported by the National Physical Science Consortium, DARPA contract DACA-76-92-C-007, LM07079-03 from the National Library of Medicine and by the New York State Science and Technology Foundation.

References

Alicia Abella. *From Imagery to Salience: Generating Locative Expressions in Context.* PhD thesis, Columbia University, 1995.

Michael Elhadad. *FUF: The Universal Unifier*, 1993.

John Fox and Nicholas Walker. Knowledge based interpretation of medical images. In *Medical Imaging*, 1989.

Berthold Horn. *Robot Vision.* The MIT Press, 1990.

Carolyn Kimme, Dana Ballard, and Jack Sklansky. Finding circles by an array of accumulators. In *Communications of the ACM*, volume 18, February 1975.

Masaharu Kobashi and Linda G. Shapiro. Knowledge-based organ identification from ct images. In *Medical Imaging VI: Image Processing*, 1992.

Yi Lu, Lucia Zamorano, Federico Moure, and Steven Schlosser. Automatic detection of boundaries of brain tumors. In *Medical Imaging VI: Image Processing*, 1992.

Peter G. Selfridge and Judith M. S. Prewitt. Organ detection in abdominal computerized tomography scans: Application to the kidney. In *Computer Graphics and Image Processing*, volume 15, pages 265–278, 1981.

H. Wechsler and J. Sklansky. Finding the rib cage in chest radiographs. In *Pattern Recognition*, 1977.